Two independent proteomic approaches provide a comprehensive analysis of the synovial fluid proteome response to Autologous Chondrocyte Implantation

Charlotte H Hulme[1,2], Emma L Wilson[2,3], Heidi R Fuller[1], Sally Roberts[1,2], James B. Richardson[1,2], Pete Gallacher[1,2], Mandy J. Peffers[4], Sally L Shirran[5], Catherine H. Botting[5], Karina T Wright[1,2]

Institutions:

1. Institute of Science and Technology in Medicine, Keele University, Keele, Staffordshire, ST5 5BG, UK

2. Robert Jones and Agnes Hunt Orthopaedic Hospital, Oswestry, Shropshire, SY10 7AG, UK

3. Chester Medical School, Chester University, Chester, CH1 4BJ, UK

4. Institute of Ageing and Chronic Disease, University of Liverpool, Liverpool, L7 8TX, UK

5. BSRC Mass Spectrometry and Proteomics Facility, University of St Andrews, North Haugh, Fife, KY16 9ST, UK

Contact Details: Charlotte H Hulme, charlotte.hulme@rjah.nhs.uk; Emma L Wilson, e.wilson@chester.ac.uk; Heidi R Fuller, h.r.fuller@keele.ac.uk; Sally Roberts, sally.roberts@rjah.nhs.uk; James B Richardson, james.richardson@rjah.nhs.uk; Peter Gallacher, peter.gallacher@rjah.nhs.uk; Mandy J. Peffers, M.J.Peffers@liverpool.ac.uk; Sally Shirran, ss101@st-andrews.ac.uk; Catherine H Blotting, cb2@st-andrews.ac.uk

Correspondence: Karina T. Wright Ph.D., ISTM, Keele University based at the RJAH Orthopaedic Hospital, Oswestry, Shropshire, UK. Telephone: +44 1691 404022; e-mail: Karina.Wright@rjah.nhs.uk

Abstract

Background: Autologous Chondrocyte Implantation (ACI) has a failure rate of approximately 20% but we are yet to fully understand why. Biomarkers are needed that can pre-operatively predict which patients are likely to fail, so that alternative or individualised therapies can be offered. We previously used a label-free (LF) quantitation with dynamic range compression proteomic approach to assess the synovial fluid (SF) of ACI responders and non-responders. However, we were only able to identify a few differentially abundant proteins at baseline. Here, we build upon these previous findings by assessing higher abundance proteins within these SFs, providing a more global proteome analysis from which we can understand more of the biology underlying ACI success or failure.

Methods: Isobaric tagging for relative and absolute quantitation (iTRAQ) proteomics was used to assess SFs from ACI responders (mean Lysholm improvement of 33; n=14) and non-responders (mean Lysholm decrease of 14; n=13) at the two stages of surgery (cartilage harvest and chondrocyte implantation). Differentially abundant proteins in iTRAQ and combined iTRAQ and LF datasets were investigated using pathway and network analyses.

Results: iTRAQ proteomics has confirmed our previous finding that there is a marked proteome shift in response to cartilage harvest (70 and 54 proteins demonstrating ≥2.0 fold change and $p<0.05$ between Stages I and II in responders and non-responders, respectively). Further, it has highlighted 28 proteins that were differentially abundant between responders and non-responders to ACI, that were not found in the LF study, 16 of which were altered at baseline. The differential expression of two proteins (complement C1S subcomponent and matrix metalloproteinase 3 (MMP3)) was confirmed biochemically. Combination of the iTRAQ and LF proteomic datasets has generated in-depth SF proteome information that has been used to generate interactome networks representing ACI success or failure. Functional pathways that are dysregulated in ACI non-responders have been identified, including acute phase response signalling.

Conclusions: Several candidate biomarkers for baseline prediction of ACI outcome have been identified. A holistic overview of the SF proteome in responders and non-responders to ACI has been profiled providing a better understanding of the biological pathways underlying clinical outcome, particularly the differential response to cartilage harvest in non-responders.

Keywords

Autologous Chondrocyte Implantation (ACI); iTRAQ proteomics; Label-free quantification proteomics; Synovial Fluid; Cartilage repair; Complement C1S subcomponent; Matrix metalloproteinase 3; MMP3

Background

Identification of putative biomarkers that can be used to predict patient outcome prior to treatment for cartilage injury has been highlighted as a key initiative for the prevention of osteoarthritis (OA) by the Osteoarthritis Research Society International (OARSI) (1). Further, the recent National Institute for Health and Care Excellence (NICE) recommendation for use of the cell therapy Autologous Chondrocyte Implantation (ACI), in the UK National Health Service (NHS) has increased the need to identify accurate prognostic biomarkers for this application (2).

We recently published the first study (3), to our knowledge, that has used a proteomic approach with the aim of identifying candidate biomarkers to predict the success of ACI, a cellular therapy for the treatment of traumatic cartilage injury (4,5). This therapy is a two-stage procedure; during the initial surgery (Stage I) healthy cartilage is harvested from a minor load-bearing region of the joint, then chondrocytes are isolated and culture expanded for three to four weeks prior to a second surgery in which the chondrocytes are implanted into the cartilage defect (Stage II) (5,6). Approximately 500 patients have been treated with ACI in our centre and despite an 81% success rate (7), we are yet to fully understand why some individuals fail to respond well. We have identified a biomarker, aggrecanase-1, that when its activity is un-detectable pre-operatively, can be used together with known demographic and injury-associated risk factors to help predict ACI success (8,9). However, we are yet to identify a biomarker (or panel of biomarkers) that can be used to accurately predict ACI failure. The identification of such a biomarker(s) for ACI and other cartilage repair strategies would allow for the better stratification of patients' prior to joint surgery and may provide candidates for therapies to improve ACI success.

Proteomic analyses remain one of the most widely used methods to identify novel biomarker candidates and have previously been utilised to identify biomarkers of OA progression (as summarised by Hsueh *et al.* in 2014 (10)). The synovial fluid (SF) provides an attractive biological fluid for biomarker identification, as it bathes the injured joint and therefore contains proteins that might reflect the whole joint environment. Proteomic profiling of the SF, however, is technically difficult due to the broad dynamic range of proteins present within it (7,8). Several un-biased, global proteomic studies for the identification of biomarkers within the SF have been completed. Nevertheless, the number of protein 'hits' has been somewhat limited, as authors have tended to

either profile SFs with no pre-treatment to account for the wide range of proteins (11–16) or have depleted high abundance proteins (17–22) meaning that the altered quantities of these proteins cannot be considered.

Isobaric tags for absolute and relative quantitation (iTRAQ) is reported as the most accurate labelling method for quantifying comparative abundance of proteins (23). When compared to label free (LF) quantitation proteomics, iTRAQ quantitation has traditionally been considered as a more accurate technique (24); however as mass spectrometers improve, these techniques are becoming more comparable and LF quantitation is becoming increasingly popular (25). Unlike LF quantitation proteomics, iTRAQ utilises isobaric tags to label the primary amines at the peptide level, prior to pooling the samples to enable simultaneous identification and quantitation of the proteins. 4plex and 8plex labels are available enabling quantitation of up to 8 conditions in a single analysis, thus minimising the number of mass spectrometry runs which can be cost effective and time efficient. However, when compared to LF quantitation, in which any number of samples can be analysed and compared, iTRAQ labelling limits the number of samples that can be compared, meaning biological replicate samples are often pooled together into relevant biological conditions. iTRAQ proteomics is a commonly used tool for the identification of biomarkers in a plethora of diseases. This proteomic approach has been used to profile the SF proteome (20,26), successfully identifying differentially abundant protein biomarker candidates for several diseases/conditions.

Our previous study highlighted the potential of using protein equalisation to study low abundance proteins in human SF, but this identified few differentially abundant proteins in baseline SF, when comparing individuals who did or did not do well following cartilage repair therapy (1). This study, therefore, aimed to increase the number of protein biomarker candidates that could be identified for the pre-operative prediction of clinical outcome following ACI and to allow for the assessment of high abundance proteins which may also strengthen our understanding of the biological processes underlying treatment success.

Methods

Synovial fluid collection and storage

SF was collected, as described previously (3,8,27), from the knee joints of patients, with informed consent and following local research ethical approval. Immediately prior to both ACI surgeries, Stage I (cartilage harvest) and Stage II (chondrocyte implantation), 20 mL of saline was injected into the joint and 20 rounds of leg flexion and extension carried out to allow aspiration of as much SF as possible (3,27). SF was then centrifuged at 6,000 *g* for 15 minutes at 4°C and split into aliquots for long-term storage in liquid nitrogen. The dilution factor of the SF samples was calculated by comparing urea content in SF to matched blood plasma using a QuantiChrom™ Urea Assay kit (BioAssay Systems, Hayward, USA) according to manufacturer's instructions, as described previously (3,8,28) and SF samples with a dilution factor greater than 10 were excluded from the study.

Clinical responders to ACI were defined as individuals who demonstrated an increase of at least 10 points in the Lysholm score at 12 months post-treatment compared to their baseline score, as has been used previously (29–31). The Lysholm score is a validated (32), patient-self assessment score, encompassing knee pain and joint function that ranges from 0-100, with 100 representing 'perfect' knee function (32,33). Thirteen patients were considered as non-responders to ACI, demonstrating a mean decrease in Lysholm score of 14 points (range -4 - -46) and 14 SF donors were considered responders with a mean improvement of 33 points (range 17-54).

Sample preparation and analysis using iTRAQ proteomics (iTRAQ nLC-MS/MS)

Total protein was quantified using a PierceTM 660nm protein assay (Thermo Scientific, Hemel Hempstead, UK) (34) and a total of 200 µg of SF protein was pooled equally from the donors in each of the following experimental groups: Stage I, responders (n=8); Stage I, non-responders (n=7); Stage II responders (n=12) and Stage II, non-responders (n=12). The pooled samples were then precipitated in six volumes of ice-cold acetone overnight at -20˚C. The precipitates were pelleted by centrifugation at 13,000 g for 10 mins at 4˚C before being re-suspended in 200 µl triethylammonium bicarbonate (TEAB) buffer. Eighty five micrograms of protein for each experimental sample was then subjected to reduction, alkylation (as instructed in the iTRAQ labelling kit (Applied Biosystems, Bleiswijk, Netherlands)). Sequencing grade modified trypsin (Promega) (10 µg per 85 µg of protein)

	Stage I		Stage II		Mann-Whitney U (p-value) (A)R v NR- SI; (B) R v NR- SI	Mann-Whitney U (p-value) (A)SI v SII-R; (B) SI v SII- NR
	Responders (n=8)	Non-responders (n=7)	Responders (n=12)	Non-responders (n=12)		
Difference in Lysholm Score	27 (17-38)	-8 (-4- -17)	34 (17-54)	-11 (-4 - -46)	(A) 0.0003; (B) <0.0001	(A) 0.21; (B) 0.55
BMI (kg/m^2)	29 (23-31)	27 (24-31)	27 (23-48)	29 (22-36)	(A) 0.94; (B) 0.54	(A) 0.73; (B) 0.68
Age (years)	32 (17-49)	40 (25-50)	40 (17-90)	43 (25-52)	(A) 0.28; (B) 0.92	(A) 0.17; (B) 0.58
Male/Female (No. of participants)	8/0	7/0	11/1	10/2	(A) >0.99; (B) >0.99	(A) >0.99; (B) 0.51
Smoker (No. of participants)	1	2	1	3	(A) 0.54; (B) 0.59	(A) >0.99; (B) >0.99
Dilution Factor of SF	5 (3-9)	4 (2-7)	4 (1-9)	3 (2-5)	(A) 0.48; (B) 0.25	(A) 0.53; (B) 0.50
Total defect area (cm^2)	14 (0.4-24)	6 (0.6-12)	6 (1-20)	5 (0.6-12)	(A) 0.74; (B) 0.35	(A) 0.45; (B) 0.28
Patella defect (No. of participants)	1	1	4	2	(A) >0.99; (B) 0.64	(A) 0.60; (B) >0.99
LFC defect (No. of participants)	2	0	0	0	(A) 0.47; (B) >0.99	(A) 0.15; (B) >0.99
LTP defect (No. of participants)	1	0	0	0	(A) >0.99; (B) >0.99	(A) 0.15; (B) >0.99
MFC defect (No. of participants)	2	2	1	6	(A) >0.99; (B) 0.07	(A) 0.54; (B) 0.63
Trochlea defect (No. of participants)	0	3	2	1	(A) 0.20; (B) >0.99	(A) 0.49; (B) 0.12
Multiple defects (No. of participants)	1	0	1	1	(A) >0.99; (B) >0.99	(A) >0.99; (B) >0.99
Unknown defect location (No. of participants)	1	1	4	2	(A) >0.99; (B) 0.64	(A) 0.60; (B) >0.99

Table 1: Demographic data for patient samples from Stage I or Stage II were analysed, those who responded well clinically (responders) or who did not respond well (non-responders) to autologous chondrocyte implantation (ACI) are indicated in separate groups. None of the demographic parameters, other than a difference in Lysholm score, showed differences between responders (R) and non-responders (NR) in individuals whose SFs from Stage I (SI) or Stage II (SII) were compared, nor were there differences between individuals who were either responders or non-responders when comparing Stage I and Stage II samples ($p \geq 0.05$; Mann-Whitney U). Data are median (range). Abbreviations: BMI, body mass index; LFC, lateral femoral condyle; LTP, lateral tibial plateau; MFC, medial femoral condyle.

was then added to the samples for overnight digestion at 37˚C. Tryptic digests were labelled with the iTRAQ tags, according to manufacturer's instructions: 114- Stage II, responders; 115- Stage II, non-responders; 116- Stage I, responders; 117- Stage I, non-responders, before being pooled to one microcentrifuge tube prior to being dried down in a vacuum centrifuge.

iTRAQ- labelled peptides were resuspended in 0.6 mL of loading Buffer Ascx (10 mM monopotassium phosphate (KH_2PO_4), 20% acetonitrile (MeCN), pH 3.0), followed by sonication. The pH was adjusted to 3.0 with 0.5 M orthophosphoric acid (H_3PO_4). The peptides were separated by strong cation exchange chromatography as described previously (35). A total of 14 SCX fractions were analysed by nanoLC ESI MSMS using a TripleTOF 5600 tandem mass spectrometer (ABSciex, Foster City, CA) as described previously (36)

The raw mass spectrometry data file was subsequently analysed using ProteinPilot 4.5 software with the Paragon™ and ProGroup™ algorithms (ABSciex) against the human sequences in the Swiss-Prot database (downloaded Dec 2012). Searches were performed using the pre-set iTRAQ settings in ProteinPilot. Trypsin was selected as the cleavage enzyme and MMTS for the modification of cysteines with a "Thorough ID" search effort. ProteinPilot's Bias correction assumes that most proteins do not change in expression. Finally, detected proteins were reported with a Protein Threshold [Unused ProtScore (confidence)] >0.05 and used in the quantitative analysis if they were identified with two or more unique peptides with 95% confidence or above. P-values and false discovery rates for the iTRAQ ratios were calculated by the ProteinPilot software. Proteins with iTRAQ ratios with p-values ≤0.05 and with differential abundances of ≥±2.0 fold change (FC) were used in further analysis.

Verification of iTRAQ nLC-MS/MS results using Enzyme Linked Immunosorbant Assay (ELISA)

Two proteins of biological relevance were measured by ELISA in the non-pooled samples to verify the mass -spectrometry findings. Firstly, Complement C1S subcomponent (C1s) was selected, as this protein demonstrated differential abundance between responders and non-responders to ACI within the baseline SF (prior to Stage I surgery) and, therefore, could have potential as a biomarker of outcome prediction. C1s was assessed using a human ELISA (Cusabio, USA). Samples were first assayed using a 1 in 100 dilution in assay sample diluent and for those samples that were undetectable in the assay was repeated using undiluted samples. Secondly Matrix metalloproteinase

3 (MMP3) was selected to investigate the differential response to Stage I surgery (i.e. the proteome shift between Stages I and II) in non-responders to ACI. MMP3 was assessed using a human Quantikine® ELISA (R&D Systems, Abingdon, UK). Samples were diluted 1 in 100 in assay kit diluent prior to assessment. Both ELISAs were carried out according to the manufacturer's instructions and protein concentrations were normalised to the sample dilution factor. Statistical analysis was performed in GraphPad Prism version 6.0. Student's t-tests were used to assess differential abundance.

Assessment of the overlap of proteins identified from the two proteomic approaches

In order to assess whether the use of two independent proteomic approaches allows for a greater number of significant protein changes to be identified, the datasets from this study (iTRAQ nLC-MS/MS) and our previously published study assessing the same patient samples (LF LC-MS/MS; (3)) were compared to one another. Venn-diagrams were plotted using VENNY 2.1.0 software (37), to assess the overlap of differentially abundant proteins that was identified via the two approaches.

Pathway and network analysis of proteomic datasets

The datasets generated from both proteomic approaches were combined. Specifically, proteins which were differentially expressed (≥1.2 FC; $p \leq 0.05$) in each biological comparison e.g. Stage I responders versus non-responders, in either proteomic approach were merged into a single dataset. A modest fold-change cut-off was used to ensure the greatest number of differentially abundant proteins could be included in the pathway and network analyses, as has been used previously (3,18). The iTRAQ nLC-MS/MS dataset independently and when merged with the LF dataset was analysed using pathway enrichment analysis (Ingenuity, Qiagen, US) to identify and visualise affected canonical pathways. Pathways with $p \leq 0.005$ were considered as statistically significant (Fisher's exact test).

The merged LF & iTRAQ nLC-MS/MS datasets of proteomic response to cartilage harvest (e.g. differential abundance between Stages I and II) in responders and non-responders were assessed using interactome network analysis, which is an unbiased mathematical method of visualising and interpreting complex interactions between large numbers of molecules (38). Interactome networks are made up of nodes (the individual objects being studied, e.g. proteins) and edges (the connections between the objects, e.g. known protein-protein interactions) (39). By studying groups of proteins that

241 are highly interconnected, known as modules, key functions within an interactome network can be
242 highlighted (39). Conducting interactome network analysis alongside pathway enrichment analysis,
243 allows for greater confidence in the selection of candidate pathways or molecules for further study as
244 these represent two independent methods of mapping the data, known protein-protein interactions
245 and text mining, respectively. The interactions between the differentially abundant proteins were
246 assessed using the Protein Interaction Network Analysis For Multiple Sets (PINA4MS) app (40) in
247 Cytoscape (v3.0) to generate network models based on protein-protein interactions. These models
248 were either based upon only those proteins identified in the proteomic analyses (non-inferred nodes)
249 or from proteins identified in the proteomic analyses alongside their inferred interactions (inferred
250 nodes) (41). The Moduland (v2.8.3) algorithm (42) was applied to the interactome networks in
251 Cytoscape (v3.0) to identify highly connected clusters of proteins (modules) that demarcate the
252 hierarchical structure of the interactome network. The biological function of each module was
253 assessed by analysing the proteins identified within each module using the pathway analysis tool in
254 Reactome software (43,44). The significance of the pathway functions identified in Reactome was
255 determined by Fisher's exact test and $p \leq 0.05$ was considered statistically significant.
256
257
258
259
260
261
262
263
264
265

Results

Proteomic data has been deposited in the PRIDE ProteomeXchange and can be accessed using the identifier PXD008321.

Identification of proteins to predict ACI outcome prior to Stage I or Stage II

iTRAQ nLC-MS/MS highlighted 16 proteins (≥±2.0 FC; p ≤0.05) which were differentially abundant between responders and non-responders to ACI at baseline (immediately prior to Stage I) (Table 2). Prior to Stage II of the ACI procedure, 12 proteins displayed differential abundance between responders and non-responders (Table 3).

At both stages of treatment, SF analysed using iTRAQ nLC-MS/MS identified a greater number of differentially abundant proteins between individuals who did or did not respond well to ACI compared to SF which had undergone protein normalisation using ProteoMiner^{TM} beads and LF LC-MS/MS analysis (3). Further, the two proteomic techniques identified no common differentially abundant proteins.

Differential abundance of proteins a Stage II compared to Stage I of ACI

Proteomic profiling of the SF using iTRAQ nLC-MS/MS highlighted a considerable effect of the cartilage harvest procedure (Stage I) in both responders and non-responders, with 70 and 54 proteins being differentially abundant between Stages I and II, respectively. Thus strengthening the similar findings from the analysis of these samples using LF LC-MS/MS (3).

Interestingly, the iTRAQ nLC-MS/MS and LF LC-MS/MS identified no common Stage I compared to Stage II protein differences in the clinical responders (70 differentially abundant proteins identified by iTRAQ nLC-MS/MS and 14 identified by LF LC-MS/MS; Table 4). This lack of overlap between the two proteomic techniques is highlighted in Figure 1. There were, however, six proteins (gelsolin, vitamin K-dependent protein S, C4b binding protein alpha chain, fibrinogen alpha chain, fibrinogen beta chain and fibrinogen gamma chain) that were identified by both proteomic techniques in the non-responders, all of which showed commonality in the direction of protein shift, across the MS platforms, with iTRAQ nLC-MS/MS consistently resulting in greater differences in abundance than those identified from the LF LC-MS/MS data. A total of 54 protein abundance changes between Stages I and II in non-responders were identified using iTRAQ nLC-MS/MS and 55 protein differences were

Protein		Fold Change	Identified using:	
Description	Accession		LF LC-MS/MS	iTRAQ nLC-MS/MS
Complement C1s subcomponent	P09871	-5.15		+
Haptoglobin	P00738	-4.49		+
Mesencephalic astrocyte-derived neurotrophic factor	P55145	2.15		+
Plasma protease C1 inhibitor	P05155	2.19		+
Ig kappa chain V-II region MIL	P01615	2.60	+	
Bifunctional glutamate/proline--tRNA ligase	P07814	2.61		+
Pigment epithelium-derived factor	P36955	3.13		+
Apolipoprotein A-IV	P06727	3.19		+
Apolipoprotein L1	O14791	3.19		+
N-acetylglucosamine-6-sulfatase	P15586	3.25		+
Retinol-binding protein 4	P02753	3.34		+
Inter-alpha-trypsin inhibitor heavy chain H1	P19827	3.37		+
Extracellular matrix protein 1	Q16610	3.77		+
Lumican	P51884	3.80		+
Histidine-rich glycoprotein	P04196	3.84		+
Endoplasmin	P14625	4.37		+
Serum paraoxonase/arylesterase 1	P27169	4.41		+

Table 2: Fold change of proteins that are differentially abundant (≥±2.0 FC; p ≤0.05; protein identified by at least 2 unique peptides) in the synovial fluid of clinical non-responders compared to clinical responders to ACI immediately prior to Stage I. Positive numbers denote higher abundance in non-responders compared to responders. Proteins were identified using either protein dynamic compression coupled with label free quantification liquid-chromatography tandem mass spectrometry (LF LC-MS/MS) or no protein dynamic compression with isobaric tags for absolute and relative quantitation (iTRAQ) LC-MS/MS.

Protein		Fold Change	Identified using:	
Description	Accession		LF LC-MS/MS	iTRAQ nLC-MS/MS
40S ribosomal protein S14	P62263	-8.63		+
Kinectin	Q86UP2	-6.20		+
Apolipoprotein C-III	P02656	-2.78		+
High mobility group protein B1	P09429	-2.56		+
Kininogen-1	P01042	2.27		+
26S protease regulatory subunit 7	P35998	2.34	+	
26S proteasome non-ATPase regulatory subunit 13	Q9UNM6	2.43	+	
Alpha-enolase	P06733	2.56		+
Alpha-2-HS-glycoprotein	P02765	2.78		+
Hemopexin	P02790	2.88		+
Ferritin light chain	P02792	2.91	+	
Platelet factor 4	P02776	3.26	+	
Thrombospondin-1	P07996	3.40	+	
Nucleosome assembly protein 1-like 1	P55209	4.94	+	
Cofilin-1	P23528	7.08	+	
EH domain-containing protein 1	Q9H4M9	7.30	+	
Hemoglobin subunit delta	P02042	8.09		+
Protein S100-A6	P06703	8.39		+
T-complex protein 1 subunit eta	Q99832	8.43	+	
Hemoglobin subunit beta	P68871	32.81		+
Hemoglobin subunit alpha	P69905	44.06		+

Table 3: Fold change of proteins that are differentially abundant (≥±2.0 FC; p ≤0.05; protein identified by at least 2 unique peptides) in the synovial fluid of clinical non-responders compared to clinical responders to ACI immediately prior to Stage II. Positive numbers denote higher abundance in non-responders compared to responders. Proteins were identified using either protein dynamic compression coupled with label free quantification liquid-chromatography tandem mass spectrometry (LF LC-MS/MS) or no protein dynamic compression with isobaric tags for absolute and relative quantitation (iTRAQ) LC-MS/MS.

Protein		Fold Change	Identified using:	
Description	**Accession**		**LF LC-MS/MS**	**iTRAQ nLC-MS/MS**
Microtubule-associated protein 1B	P46821	-20.65	+	
40S ribosomal protein S14	P62263	-16.75		+
Protein disulfide-isomerase A6	Q15084	-7.59		+
Nucleolin	P19338	-5.11		+
Histone H1.2	P16403	-3.84		+
Stress-induced-phosphoprotein 1	P31948	-3.63		+
Complement factor D	P00746	-3.44		+
SH3 domain-binding glutamic acid-rich-like protein	O75368	-3.44		+
Heterogeneous nuclear ribonucleoprotein U	Q00839	-3.40		+
78 kDa glucose-regulated protein	P11021	-3.25		+
Cartilage oligomeric matrix protein	P49747	-3.10		+
Annexin A2	P07335	-2.96		+
Mesencephalic astrocyte-derived neurotrophic factor	P55145	-2.86		+
Kinectin	Q86UP2	-2.81		+
Complement factor H-related protein 3	Q02985	-2.77	+	
Phosphatidylethanolamine-binding protein 1	P30086	-2.51		+
Peroxiredoxin-4	Q13162	-2.49	+	
Regucalcin	Q15493	-2.44		+
Malate dehydrogenase, mitochondrial	P40926	-2.44		+
N-acetylglucosamine-6-sulfatase	P15586	-2.31		+
Gelsolin	P06396	-2.27		+
Alpha-endosulfine	O43768	-2.25		+
Peptidyl-prolyl cis-trans isomerase FKBP3	Q00688	-2.11		+
Hemopexin	P02790	2.05		+
Serum paraoxonase/arylesterase 1	P27169	2.07		+
Secreted phosphoprotein 24	Q13103	2.10	+	
Heparin cofactor 2	P05546	2.13		+
Ferritin light chain	P02792	2.21	+	
Attractin	O75882	2.21		+
Ig gamma-2 chain C region	P01859	2.23		+
Plasma kallikrein	P03952	2.24		+
Chondroitin sulfate proteoglycan 4	Q6UVK1	2.35	+	
Collagen alpha-2(I) chain	P08123	2.37	+	
Collagen alpha-1(V) chain	P20908	2.54	+	
CD5 antigen-like	O43866	2.58		+
Phospholipid transfer protein	P55058	2.63		+
Insulin-like growth factor-binding protein complex acid labile subunit	P35858	2.68		+
Prothrombin	P00734	2.68		+
Beta-2-glycoprotein 1	P02749	2.78		+
Collagen alpha-2(V) chain	P05997	2.84	+	
Plasma protease C1 inhibitor	P05155	2.91		+

Serum amyloid P-component	P02743	2.91		+
Complement C1q subcomponent subunit B	P02746	3.01		+
Collagen alpha-1(I) chain	P02452	3.05	+	
Alpha-2-antiplasmin	P08697	3.10		+
Alpha-1B-glycoprotein	P04217	3.19		+
Complement factor B	P00751	3.25		+
Complement component C7	P10643	3.40		+
Vitamin K-dependent protein S	P07225	3.42		+
Apolipoprotein E	P02649	3.44		+
Alpha-1-antichymotrypsin	P01011	3.44		+
Carboxypeptidase N subunit 2	P22792	3.53		+
Vitronectin	P04004	3.63		+
Inter-alpha-trypsin inhibitor heavy chain H3	Q06033	3.66		+
Complement C5 O	P01031	4.00		+
Plasminogen	P00747	4.06		+
Kininogen-1	P01042	4.17		+
Platelet factor 4	P02776	4.26	+	
Inter-alpha-trypsin inhibitor heavy chain H2	P19823	4.49		+
Periostin	Q15063	4.57	+	
Apolipoprotein L1	O14791	4.61		+
Protein 4.1	P11171	4.66	+	
26S proteasome non-ATPase regulatory subunit 13	Q9UNM6	4.78	+	
Inter-alpha-trypsin inhibitor heavy chain H1	P19827	5.01		+
Inter-alpha-trypsin inhibitor heavy chain H4	Q14624	5.06		+
Complement C1r subcomponent	P00736	5.15		+
Complement component C6	P13671	5.45		+
Complement factor H	P08603	5.50		+
Catalase	P04040	5.60		+
Ficolin-3	O75636	6.43		+
C4b-binding protein alpha chain	P04003	7.05		+
Ceruloplasmin	P00450	7.51		+
Pregnancy zone protein	P20742	8.09		+
Fibrinogen alpha chain	P02671	8.40		+
Apolipoprotein M	O95445	9.04		+
Protein S100-A6	P06703	9.82		+
Hemoglobin subunit alpha	P69905	9.82		+
Complement C1s subcomponent	P09871	10.00		+
Ig mu chain C region	P01871	12.13		+
Haptoglobin	P00738	13.68		+
Fibrinogen beta chain	P02675	16.90		+
Hemoglobin subunit beta	P68871	19.41		+
Fibrinogen gamma chain	P02679	23.55		+

328 **Table 4**: Fold change of proteins that are differentially abundant (≥±2.0 FC; p ≤0.05; protein identified
329 by at least 2 unique peptides) in the synovial fluid of clinical responders at Stage II compared to Stage

I of ACI. Positive numbers denote higher abundance at Stage II compared to Stage I. Proteins were identified using either protein dynamic compression coupled with label free quantification liquid-chromatography tandem mass spectrometry (LF LC-MS/MS) or no protein dynamic compression with isobaric tags for absolute and relative quantitation (iTRAQ) LC-MS/MS.

identified by LF LC-MS/MS (Table 5; Figure 1).

iTRAQ nLC-MS/MS confirmed that there is a significant response to the cartilage harvest procedure (Stage I) in non-responders to ACI

Pathway analysis of the iTRAQ nLC-MS/MS identified proteins, using the pathway enrichment tools in Ingenuity, suggested that the proteins which were differentially abundant at Stage II compared to Stage I in non-responders are likely to impact on numerous canonical pathways; many of which were confirmatory of the previously published functional pathways identified from the LF nLC-MS/MS derived proteins (3). These functional pathways included acute phase response signalling ($p=2.93 \times 10^{-1}$), the complement system ($p=2.11 \times 10^{-1}$) and Liver X receptor/ Retinoic X receptor (LXR/RXR) signalling ($p=1.95 \times 10^{-1}$). Moreover, many more functional pathways were affected as a result of the proteins that were differentially abundant in response to Stage II compared to Stage I in non-responders compared to responders (Supplementary Tables 1 and 2); reiterating that the SF proteome response to cartilage harvest is more distinct in non-responders to ACI.

Similar pathways were identified from the differentially abundant proteins identified by the iTRAQ nLC-MS/MS and LF LC-MS/MS analyses

Both iTRAQ nLC-MS/MS and LF LC-MS/MS analyses resulted in the acute phase response signalling being highlighted as one of the most significantly affected pathways in response to cartilage harvest in non-responders to ACI, therefore this pathway was further assessed. Figure 4 highlights that analysis of the SF proteome using the two independent proteomic techniques resulted in a greater number of differentially abundant downstream proteins being identified. In addition, many complementary proteins have been identified when comparing these datasets, with the vast majority of proteins that are predicted to be increased in the plasma (the standard bodily fluid referred to in Ingenuity) during the acute phase response being more abundant in the SF at Stage II compared to Stage I and vice versa.

As the results of the two proteomic approaches seem to be complementary to one another, the two datasets were combined to generate a more comprehensive profile of the SF proteome. Pathway analysis using Ingenuity again identified many similar functional pathways as were identified via the independent LF LC-MS/MS and iTRAQ nLC-MS/MS datasets. The most significant canonical

Protein		Fold Change	Identified using:	
Description	Accession		LF LC-MS/MS	iTRAQ nLC-MS/MS
Protein S100-A6	P06703	-4.49		+
Annexin A1	P04083	-4.13		+
Hemoglobin subunit beta	P68871	-4.09		+
Complement factor D	P00746	-3.87		+
Perilipin-4	Q96Q06	-3.87	+	
Gelsolin	*P06396*	-3.31		+
Gelsolin	*P06396*	-1.68	+	
Syntaxin-7	O15400	-3.31	+	
Fermitin family homolog 3	Q86UX7	-3.29	+	
Histone H1.2	P16403	-3.13		+
Transaldolase	P37837	-3.08		+
Neuroblast differentiation-associated protein AHNAK	Q09666	-2.78	+	
Heterogeneous nuclear ribonucleoprotein K	P61978	-2.69	+	
Hyaluronan and proteoglycan link protein 3	Q96S86	-2.65	+	
Alpha-enolase	P06733	-2.63		+
ATP-citrate synthase	P53396	-2.63	+	
Annexin A2	P07355	-2.56		+
Fatty acid-binding protein, epidermal	Q01469	-2.43	+	
Peroxiredoxin-1	Q06830	-2.20	+	
Tripeptidyl-peptidase 1	O14773	-2.19	+	
Insulin-like growth factor-binding protein 6	P24592	-2.13	+	
Na(+)/H(+) exchange regulatory cofactor NHE-RF1	O14745	-2.11	+	
Peroxiredoxin-6	P30041	-2.08	+	
Histamine N-methyltransferase	P50135	-2.07	+	
Mortality factor 4-like protein 1	Q9UBU8	-2.06	+	
Transcription elongation factor A protein 1	P23193	-2.06	+	
Cartilage acidic protein 1	Q9NQ79	-2.03		+
2',3'-cyclic-nucleotide 3'-phosphodiesterase	P09543	-1.20	+	
Fructose-bisphosphate aldolase A	P04075	-1.97	+	
Leucine zipper transcription factor-like protein 1	Q9NQ48	-1.94	+	
Protein S100-A13	Q99584	-1.94	+	
40S ribosomal protein S3	P23396	-1.93	+	
Filamin-A	P21333	-1.92	+	
Microtubule-associated protein RP/EB family member 1	Q15691	-1.92	+	
Nuclear migration protein nudC	Q9Y266	-1.90	+	
Prostaglandin E synthase 3	Q15185	-1.85	+	
Stress-induced-phosphoprotein 1	P31948	-1.85	+	
Cytokine-like protein 1	Q9NRR1	-1.81	+	
Plastin-2	P13796	-1.81	+	
Coronin-1C	Q9ULV4	-1.80	+	
Vinculin	P18206	-1.80	+	

Cathepsin K	P43235	-1.79	+	
Hsc70-interacting protein	P50502;Q8IZP2	-1.76	+	
Putative phospholipase B-like 2	Q8NHP8	-1.74	+	
Spectrin beta chain, erythrocytic	P11277	-1.73	+	
Complement factor I	P05156	2.11		+
Alpha-1-antichymotrypsin	P01011	2.22		+
Titin	Q8WZ42	2.23		+
Cytoplasmic dynein 1 heavy chain 1	Q14204	2.23	+	
F-actin-capping protein subunit beta	P47756	2.25	+	
Mannan-binding lectin serine protease 1	P48740	2.26	+	
Serum amyloid P-component	P02743	2.27		+
Complement component C6	P13671	2.29		+
Thrombospondin-3	P49746	2.36	+	
Soluble scavenger receptor cysteine-rich domain-containing protein SSC5D	A1L4H1	2.39	+	
Plasma kallikrein	P03952	2.42		+
Complement factor B	P00751	2.47		+
Afamin	P43652	2.47		+
Vitamin K-dependent protein S	P07225	2.49	+	
Vitamin K-dependent protein S	P07225	3.08		+
Integrin beta-like protein 1	O95965	2.51	+	
C4b-binding protein beta chain	P20851	2.55	+	
Fibronectin	P02751	2.58	+	
Clusterin	P10909	2.65		+
Vitronectin	P04004	2.68		+
Bifunctional glutamate/proline--tRNA ligase	P07814	2.70		+
Nucleobindin-1	Q02818	2.71	+	
Complement component C9	P02748	2.75		+
Zinc-alpha-2-glycoprotein	P25311	2.75		+
Complement C1r subcomponent	P00736	2.83		+
Heparin cofactor 2	P05546	2.83		+
Ferritin light chain	P02792	2.84	+	
Proteoglycan 4	Q92954	2.88		+
C4b-binding protein alpha chain	P04003	2.91	+	
C4b-binding protein alpha chain	P04003	10.38		+
Matrix Metalloproteinase 3	P08254	2.91		+
Attractin	O75882	2.94		+
Insulin-like growth factor-binding protein complex acid labile subunit	P35858	3.02		+
Alpha-1B-glycoprotein	P04217	3.05		+
Fibrinogen alpha chain	P02671	3.10	+	
Fibrinogen alpha chain	P02671	11.91		+
Lumican	P51884	3.13		+
Chondroitin sulfate proteoglycan 4	Q6UVK1	3.16	+	
Collagen alpha-2(V) chain	P05997	3.19	+	

Complement C2	P06681	3.22		+
Fibrinogen beta chain	P02675	3.25	+	
Fibrinogen beta chain	P02675	18.37		+
Secreted phosphoprotein 24	Q13103	3.26	+	
Matrix Metalloproteinase 1	P03956	3.33	+	
Latent-transforming growth factor beta-binding protein 1	Q14766	3.45	+	
Phospholipid transfer protein	P55058	3.47		+
Inter-alpha-trypsin inhibitor heavy chain H3	Q06033	3.47		+
Complement C1q tumor necrosis factor-related protein 3	Q9BXJ4	3.50	+	
Adipocyte enhancer-binding protein 1	Q8IUX7;Q8N436	3.51	+	
Adiponectin	Q15848	3.52	+	
Fibrinogen gamma chain	P02679	3.79	+	
Fibrinogen gamma chain	P02679	18.37		+
Plasminogen	P00747	3.84		+
Apolipoprotein C-II	P02655	3.94		+
CD5 antigen-like	O43866	4.17		+
Collagen alpha-1(V) chain	P20908	4.26	+	
Complement factor H	P08603	4.57		+
Inter-alpha-trypsin inhibitor heavy chain H4	Q14624	4.61		+
Complement C5	P01031	4.74		+
Collagen alpha-1(I) chain	P02452	4.84	+	
Ceruloplasmin	P00450	5.01		+
Histidine-rich glycoprotein	P04196	5.11		+
Target of Nesh-SH3	Q7Z7G0	5.25		+
Plasma protease C1 inhibitor	P05155	5.30		+
Apolipoprotein M	O95445	5.65		+
Inter-alpha-trypsin inhibitor heavy chain H2	P19823	5.70		+
Periostin	Q15063	5.81	+	
Apolipoprotein C-III	P02656	5.92		+
Kinectin	Q86UP2	6.02		+
Carboxypeptidase N subunit 2	P22792	7.73		+
Serum paraoxonase/arylesterase 1	P27169	8.02		+
Apolipoprotein L1	O14791	8.95		+
Ig mu chain C region	P01871	13.30		+
Apolipoprotein E	P02649	13.80		+
Inter-alpha-trypsin inhibitor heavy chain H1	P19827	15.56		+

Table 5: Fold change of proteins that are differentially abundant (≥±2.0 FC; p ≤0.05; protein identified by at least 2 unique peptides) in the synovial fluid of clinical non-responders at Stage II compared to Stage I. Positive numbers denote higher abundance at Stage II compared to Stage I of ACI. Proteins were identified using either protein dynamic compression coupled with label free quantification liquid-chromatography tandem mass spectrometry (LF LC-MS/MS) or no protein dynamic compression with isobaric tags for absolute and relative quantitation (iTRAQ) LC-MS/MS. Proteins identified by both proteomic techniques are underlined and in italics.

pathways associated with the non-responder response to cartilage harvest (Stage II vs Stage I) were, acute phase response signalling ($p=1.10 \times 10^{-9}$), intrinsic prothrombin activation pathway ($p= 3.43X10^{-7}$) and the complement system ($p=1.22x10^{-6}$). Further, analysis of upstream regulators to these dysregulated proteins included those identified using the LF LC-MS/MS analysis data alone e.g. transforming growth factor beta 1 (TGFB1; $p=2.05 \times 10^{-13}$), dihydrotestosterone ($p= 4.48 \times 10^{-11}$) and peroxisome proliferator-activated receptor alpha (PPARA; $p=1.09 \times 10^{-9}$) (3).

The combined datasets were then used to generate unbiased interactome networks that represent the differentially abundant proteins (non-inferred networks), their likely interacting proteins (inferred networks) and how these proteins interact with one another, resulting in models of systemic protein response to cartilage harvest in either the responders or non-responders to ACI. Based on proteins that were differentially abundant between Stages I and II of ACI in non-responders, an interactome network consisting of 115 nodes (proteins) and 40 edges (protein-protein interactions) was generated. Further, an inferred network consisting of 2893 proteins and 35576 protein-protein interactions was generated based upon the addition of proteins that are likely to interact with the differentially abundant proteins (PINA4MS interactome database). Proteins that were differentially abundant in response to cartilage harvest in responders to ACI were used to generate interactome networks (non-inferred: 83 nodes and 118 edges; inferred: 2084 nodes and 54007 edges). The ModuLand algorithm was applied to each of these networks to identify modules within the network that can be hierarchically ranked to identify groups of proteins that are the most fundamental in the functioning of the network. Figure 4 highlights the top 10 modules from each of the networks generated. These modules again highlight the disparity between the ACI responder and non-responder response to cartilage harvest, with only modules centred on the proto-oncogene tyrosine-protein kinase (SRC) protein being identified in the inferred networks of both non-responder and responder groups. Interestingly, assessment of the functional pathways related to the ModuLand identified modules in the non-responder networks again highlighted regulation of the complement cascade ($p= 1.68x10^{-8}$; Fisher's Exact test), thus providing confidence in its importance based on identification via two independent bioinformatic approaches.

Discussion

The recent NICE technology appraisal of ACI has recommended this treatment for a specific subset of patients with cartilage injury in the knee (2). The identification of novel biomarkers that can strengthen current patient demographic risk factors in predicting clinical outcomes (9), as well as developing a greater understanding of the underlying biology associated with success and failure will be beneficial, particularly as this treatment option is likely to be implemented on a wider scale in the near future. This study has built upon our previously published work (3,8), highlighting a number of novel protein candidates that have potential as biomarkers to predict ACI outcome. Moreover, comprehensive proteomic profiling of SF has further highlighted proteome differences between responders and non-responders to ACI.

In the majority of studies in which the SF proteome has been profiled, either high (11–16) or low (17–22) abundance proteins have been assessed via depletion or non-depletion of abundant proteins prior to proteomic analysis. Our study highlights that the use of both a proteome dynamic range compression technique (Proteominer^{TM}) (3) in tandem with analysis of non-depleted SF samples can provide a more holistic overview of proteome changes; since both iTRAQ nLC-MS/MS and LF LC-MS/MS highlighted large numbers of differentially abundant proteins between Stages I and II of ACI, with little crossover between techniques. This type of all-inclusive approach to unbiased whole-proteome analysis of biological fluids may therefore be more successful in the identification of candidate biomarkers for other treatments/ disease states beyond those investigated here.

A limitation of our previous study (3) was that very few proteins were identified as differentially abundant between responders and non-responders at baseline. In order for biomarkers aimed at predicting ACI success to be most useful clinically, patients who are likely to fail or respond to this procedure need to be identified prior to any surgical intervention. Interestingly, analysis of non-dynamic range compressed proteins with iTRAQ nLC-MS/MS analysis was able to detect a greater number of differentially abundant proteins between responders and non-responders prior to Stage I surgery. The protein with most altered abundance in responders compared to non-responders at Stage I was C1s. This higher abundance in responders was confirmed in individual patient samples using a biochemical assay. C1s is a major constituent of the trimeric complement C1 protein, which triggers the classical complement pathway. Once activated, the classical complement pathway

promotes inflammation to enable the removal of damaged cells and/or microbes. Moreover, C1s has been shown to cleave insulin growth factor 1 (IGF-1) (45) and insulin like growth factor binding protein 5 (IGFBP-5) (46). Both IGF-1 and IGFBP-5 are chondroprotective when in their intact state (45,47) and inhibition of C1s activity within the canine SF reduced cleavage of IGFBP-5 and IGF-1, resulting in reduced cartilage damage following anterior cruciate ligament rupture (45). These studies indicate that high C1s activity levels are likely detrimental to cartilage repair. Further, the complement cascade is known to be important in the pathogenesis of OA, with OA patients demonstrating increased gene expression of complement agonists compared to inhibitors (48). OA related pathogenesis, such as the release of cartilage extracellular matrix molecules and the production of inflammatory mediators all induce complement activation (48). The increased pre-operative levels that we have identified in individuals who responded well to ACI, perhaps indicate that ACI has potential to be successful in individuals who may have developed an early OA phenotype.

Analysis of the iTRAQ nLC-MS/MS and LF LC-MS/MS datasets, both independently and when combined, has highlighted that there is a marked proteome shift in response to cartilage harvest, i.e. between Stages I and II of ACI. This analysis has resulted in a plethora of candidate biomarkers that may have the potential to inform as to whether an individual is likely to respond well to ACI prior to chondrocytes being implanted during Stage II. The proteoglycan, collagen II, IX and X degrading enzyme, MMP3 (49) has been biochemically validated as one of these candidate proteins that is significantly increased at Stage II compared to Stage I only in non-responders to ACI. Use of these biomarkers could have the potential to prevent the burden of a second surgery in a patient for whom this therapy is likely to be unsuccessful, could indicate that a greater period of time should be left from when the cartilage harvest procedure takes place and when the cells are implanted or that a tailored cartilage implantation procedure would be more efficacious.

To investigate the significant proteome shift that exists in response to cartilage harvest, pathway analyses were performed to better distinguish the underlying biological mechanisms which dictate if an individual will respond to ACI or not. The acute phase response was the pathway predicted to be most significantly differentially regulated in response to cartilage harvest in non-responders to ACI. In-depth assessment of individual protein changes within this pathway again highlighted the benefit of

using independent proteomic techniques to profile the SF, as a large number of proteins were differentially abundant between Stages I and II, only three of which were identified using both techniques. The acute phase response is the body's first systemic response to immunological stress, trauma and surgery (52). At the site of injury/trauma, pro-inflammatory cytokines are normally released, activating inflammatory cells, ultimately resulting in inflammatory mediators and cytokines being released into the extracellular fluid compartment to be circulated within the blood (52). Interestingly, previous bioinformatic analyses of the proteome of late OA compared to healthy controls, highlighted dysregulated acute phase response in the end-stage OA cohort (18). The exacerbated activation of the acute phase response in non-responders following initial surgery could indicate that these patients have a greater immune response to surgery and that they have a lesser ability to dampen down the acute phase following surgery or that they have already developed an advanced OA phenotype, deeming a therapy to repair cartilage injury unsuitable.

Finally, the datasets of combined iTRAQ nLC-MS/MS and LF LC-MS/MS identified proteins were used to generate interactome models that represent the systemic proteome response to cartilage harvest that exists within the SF of both ACI responders and non-responders, from which biological functional pathways could be further studied. Biological functional pathways that were identified using this approach, as well as using IPA can most confidently be taken forward as candidates for further study, as they have been identified by independent bioinformatic methods. Furthermore, given the complexity of the knee joint environment, it is likely that the responder/non-responder phenotype is the result of many subtle protein changes which together contribute to overall dysfunction of a biological network, rather than the result of an individual biological molecule or pathway *per se.* Therefore, the interactome models generated in this study provide an important opportunity to consider how these proteins interact with one another and result in such phenotypes, as well as, providing a platform for further studies to investigate how potential modifications to the ACI procedure e.g. using co-incidental anti-inflammatory drugs in non-responders at Stage II, may alter these biological networks. Thus, these models may provide a potential *in silico* tool for predicting ACI outcome, as is commonly used in drug development strategies (53).

Conclusion

This study has highlighted the advantage of using two independent proteomic techniques to profile a holistic overview of the SF proteome, ideal for unbiased identification of biomarker candidates. iTRAQ nLC-MS/MS analysis of SF samples from individuals who have either responded well or very poorly to ACI has highlighted proteins that with further validation have the potential to predict clinical outcome prior to treatment. We have confirmed that there is a marked SF proteome shift following cartilage injury, which is exacerbated in non-responders. Network and pathway analyses have demonstrated the complexity of the biological response underlying this proteome shift in non-responders, with several biological pathways identified that may act as targets for therapeutic intervention.

List of Abbreviations

ACI, autologous chondrocyte implantation; BMI, body mass index; C1s, Complement 1S subcomponent; iTRAQ, isobaric tagging for relative and absolute quantitation; LC-MS/MS, liquid chromatography- tandem mass spectrometry; LF, label-free quantitation; LFC, lateral femoral condyle; LTP, lateral tibial plateau; MFC, medial femoral condyle; LXR/RXR, Liver X receptor/ Retinoic X receptor; MMP3, matrix metalloproteinase 3; nLC-MS/MS, nano liquid chromatography-tandem mass spectrometry; NICE, The National Institute for Health and Care Excellence; NHS, National Health Service OA, osteoarthritis; OARSI, Osteoarthritis Research Society International; PINA4MS, Protein Interaction Network Analysis For Multiple Sets; SF, synovial fluid; SRC, proto-oncogene tyrosine-protein kinase; TEAB, triethylammonium bicarbonate; TGFB1, transforming growth factor beta 1

Declarations

Ethics approval and consent to participate

SF samples from patients undergoing ACI were collected under three independent ethical approvals: ‘Investigating the potential for cells and molecules isolated from orthopaedic patients for modelling and understanding pathogenic conditions and developing diagnostic markers and therapies for musculoskeletal disorders and spinal cord injury’ (11/NW/0875); ‘Autologous cell therapy for Osteoarthritis: An evaluation of the safety and efficacy of autologous transplantation of articular chondrocytes and/or bone marrow derived stromal cells to repair chondral/osteochondral lesions of the knee’ (11/WM/0175) and ‘Arthritis and cartilage repair study’ (06/Q6201/9). 11/NW/0875 was

approved by the NRES committee North West- Liverpool East. 11/WM/0175 was approved by the NRES committee West Midlands - Coventry and Warwick and 06/Q2601/9 was approved by Shropshire and Staffordshire-Shropshire local research ethics committee. All patients gave valid informed consent prior to samples being collected.

Consent for publication

Not applicable for this study.

Availability of data and material

Proteomic data has been deposited in the PRIDE ProteomeXchange and can be accessed using the identifier PXD008321.

Competing interests

The authors declare that they have no competing interests.

Funding

We would like to thank Arthritis Research UK for supporting this work via grants 19429, 20815 and 21122. The sponsors had no involvement in the study design, data collection and interpretation or preparation of the manuscript. Mandy Peffers is supported through a Wellcome Trust Clinical Intermediate Fellowship. This work was supported by the Wellcome Trust grant 094476/Z/10/Z which funded the purchase of the TripleTOF 5600 mass spectrometer at the BSRC Mass Spectrometry and Proteomics Facility, University of St Andrews.

Authors' contributions

CHH, ELW, KTW, HRF, MJP & SR came up with conception and design of the study. CHH, ELW, HRJ, SLS & CHB collected data which was then analysed and interpreted by CHH, ELW, KTW, MJP & HRF. CHH, ELW, HRF, SR, MJP, SLS, CHB, JBR, PG & KTW drafted the manuscript, critically revised and approved the final article. PG & JBR provided patients' synovial fluid samples. Funding for the study was obtained by KTW & SR. All authors read and approved the final manuscript.

Acknowledgments

We would like to thank Professor Rob Beynon for his advice on study design and help with analysis of samples using label free quantitation proteomic analysis.

References

1. Kraus VB, Blanco FJ, Englund M, Henrotin Y, Lohmander LS, Losina E, et al. OARSI Clinical Trials Recommendations: Soluble biomarker assessments in clinical trials in osteoarthritis. Osteoarthr Cartil. 2015;23(5):686–97.

2. National Institue for Health and Care Excellence. Autologous chondrocyte implantation for treating symptomatic articular cartilage defects of the knee [Internet]. 2017. Available from: https://www.nice.org.uk/guidance/ta477/resources/autologous-chondrocyte-implantation-for-treating-symptomatic-articular-cartilage-defects-of-the-knee-pdf-82604971061701

3. Hulme CH, Wilson EL, Peffers MJ, Roberts S, Simpson DM, Richardson JB, et al. Autologous chondrocyte implantation-derived synovial fluids display distinct responder and non-responder profiles. Arthritis Res Ther. 2017;19:150.

4. Gillogly SD, Voight M, Blackburn T. Treatment of articular cartilage defects of the knee with autologous chondrocyte implantation. J Orthop Sport Phys Ther. 1998;28(4):241–51.

5. Richardson JB, Caterson B, Evans EH, Ashton BA, Roberts S. Repair of human articular cartilage after implantation of autologous chondrocytes. J Bone Joint Surg Br. 1999;81(6):1064–8.

6. Wright KT, Mennan C, Fox H, Richardson JB, Banerjee R, Roberts S. Characterization of the cells in repair tissue following autologous chondrocyte implantation in mankind: a novel report of two cases. Regen Med. 2013;8:699–709.

7. Bhosale AM, Kuiper JH, Johnson WE, Harrison PE, Richardson JB. Midterm to long-term longitudinal outcome of autologous chondrocyte implantation in the knee joint: a multilevel analysis. Am J Sports Med. 2009;37(Suppl 1):131S – 8S.

8. Wright KT, Kuiper JH, Richardson JB, Gallacher P, Roberts S. The absence of detectable ADAMTS-4 (aggrecanase-1) activity in synovial fluid is a predictive indicator of autologous chondrocyte implantation success. Am J Sports Med. 2017;45(8):1806–14.

9. Dugard MN, Herman Kuiper J, Parker J, Roberts S, Robinson E, Harrison P, et al. Development of a Tool to Predict Outcome of Autologous Chondrocyte Implantation. Cartilage. 2016;1–12.

10. Hsueh MF, Onnerfjord P, Kraus VB. Biomarkers and proteomic analysis of osteoarthritis. Matrix Biol. Elsevier B.V.; 2014;39:56–66.

11. Chiaradia E, Pepe M, Tartaglia M, Scoppetta F, D'Ambrosio C, Renzone G, et al. Gambling on putative biomarkers of osteoarthritis and osteochondrosis by equine synovial fluid proteomics. J Proteomics. 2012;75:4478–93.

12. Pan X, Huang L, Chen J, Dai Y, Chen X. Analysis of synovial fluid in knee joint of osteoarthritis: 5 Proteome patterns of joint inflammation based on matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. Int Orthop. 2012;36(1):57–64.

13. Sohn DH, Sokolove J, Sharpe O, Erhart JC, Chandra PE, Lahey LJ, et al. Plasma proteins present in osteoarthritic synovial fluid can stimulate production via Toll-like receptor 4. Arthritis Res Ther. 2012;14(1):R7.

14. Liao W, Li Z, Wang H, Wang J, Fu Y, Bai X. Proteomic analysis of synovial fluid: Insight into the pathogenesis of knee osteoarthritis. Int Orthop. 2013;37(6):1045–53.

15. Noh R, Park SG, Ju JH, Chi SW, Kim S, Lee CK, et al. Comparative proteomic analyses of

synovial fluids and serums from rheumatoid arthritis patients. J Microbiol Biotechnol. 2014;24(1):119–26.

16. Liao W, Li Z, Zhang H, Li J, Wang K, Yang Y. Proteomic analysis of synovial fluid as an analytical tool to detect candidate biomarkers for knee osteoarthritis. Int J Clin Exp Pathol. 2015;8(9):9975–89.

17. Mateos J, Lourido L, Fernández-Puente, P Calamia V, Fernández-López, C Oreiro N, Ruiz-Romero C, Blanco FJ. Differential protein profiling of synovial fluid from rheumatoid arthritis and osteoarthritis patients using LC–MALDI TOF/TOF. J Proteomics. 2012;75(10):2869–78.

18. Ritter SY, Subbaiah R, Bebek G, Crish J, Scanzello CR, Krastins B, et al. Proteomic analysis of synovial fluid from the osteoarthritic knee: Comparison with transcriptome analyses of joint tissues. Arthritis Rheum. 2013;65(4):981–92.

19. Balakrishnan L, Nirujogi RS, Ahmad S, Bhattacharjee M, Manda SS, Renuse S, et al. Proteomic analysis of human osteoarthritis synovial fluid. Clin Proteomics. 2014;11(1):1–13.

20. Balakrishnan L, Bhattacharjee M, Ahmad S, Nirujogi RS, Renuse S, Subbannayya Y, et al. Differential proteomic analysis of synovial fluid from rheumatoid arthritis and osteoarthritis patients. Clin Proteomics. 2014;11(1):1.

21. Bennike T, Ayturk U, Haslauer CM, Froehlich JW, Proffen BL, Barnaby O, et al. A normative study of the synovial fluid proteome from healthy porcine knee joints. J Proteome Res. 2014;13(10):4377–87.

22. Bhattacharjee M, Balakrishnan L, Renuse S, Advani J, Goel R, Sathe G, et al. Synovial fluid proteome in rheumatoid arthritis. Clin Proteomics. BioMed Central; 2016;1–11.

23. Wu WW, Wang G, Baek SJ, Shen RF. Comparative study of three proteomic quantitative methods, DIGE, cICAT, and iTRAQ, using 2D gel- or LC-MALDI TOF/TOF. J Proteome Res. 2006;5(3):651–8.

24. Bantscheff M, Schirle M, Sweetman G, Rick J, Kuster B. Quantitative mass spectrometry in proteomics: a critical review. Anal Bioanal Chem. 2007;389(4):1017–31.

25. Lilley KS, Beynon RJ, Eyers CE, Hubbard SJ. Focus of Quantitative Proteomics. Proteomics. 2015;15(18):3101–3.

26. Herr MM, Fries KM, Upton LG, Edsberg LE. Potential biomarkers of temporomandibular joint disorders. J Oral Maxillofac Surg. 2011;69(1):41–7.

27. Roberts S, Evans H, Wright K, van Niekerk L, Caterson B, Richardson JB, et al. ADAMTS-4 activity in synovial fluid as a biomarker of inflammation and effusion. Osteoarthritis Cartilage. 2015;23(9):1622–6.

28. Kraus V, Stabler T, Kong S, Varjum G, McDaniel G. Measurement of synovial fluid volume using urea. Osteoarthr Cartil. 2007;15(1217):e20.

29. Ehrich E, Davies G, Watson D, Bolohnese J, Seidenberg B, Bellamy N. Minimal perceptible clinical improvement with the Western Ontario and McMaster Universities osteoarthritis index questionnaire and global assessments in patients with osteoarthritis. J Rheumatol. 2000;27(11):2635–41.

30. Roos E, Lohmander L. The Knee Injury and Osteoarthritis Outcome Score (KOOS): from joint injury to osteoarthritis. Health Qual Life Outcomes. 2003;1:64.

31. Saris D, Vanlauwe J, Victor J, Almqvist K, Verdonk R, Bellemans J, et al. Treatment of symptomatic cartilage defects in the knee: characterized chondrocyte implantation results in better clinical outcome at 36 months in a randomized trial compared to microfracture. Am J Sports Med. 2009;37(Suppl 1):10S – 19S.

32. Smith H, Richardson J, Tennant A. Modification and validation of the Lysholm Knee Scale to assess articular cartilage damage. Osteoarthr Cartil. 2009;17:55–8.

33. Lysholm J, Gillquist J. Evaluation of knee ligament surgery results with special emphasis on use of a scoring scale. Am J Sports Med. 1982;10(3):150–4.

34. Stoscheck CM. Protein assay sensitive at nanogram levels. Anal Biochem. 1987;160(2):301–5.

35. Fuller HR, Mandefro B, Shirran SL, Gross AR, Kaus A, Botting CH, et al. Spinal muscular atrophy patient iPSC-derived motor neurons have reduced expression of proteins important in neuronal development. Front Cell Neurosci. 2016;9(506).

36. Fuller H, Slade R, Jovanov-Milošević N, Babić M, Sedmak G, Šimić G, et al. Stathmin is enriched in the developing corticospinal tract. Mol Cell Neurosci. 2015;12(69):12–21.

37. Oliveros JC. VENNY. An interactive tool for comparing lists with Venn Diagrams. BioinfoGP of CNB-CSIC. 2007. p. http://bioinfogp.cnnb.csic.es/tools/venny/index.ht.

38. Albert R, Barabasi AL. Statistical mechanics of complex networks. Rev Mod Phys. 2002;74(1):47–97.

39. Vidal M, Cusick ME, Barabási A-L. Interactome networks and human disease. Cell. Cell Press; 2011;144(6):986–98.

40. Cowley MJ, Pinese M, Kassahn KS, Waddell N, Pearson J V., Grimmond SM, et al. PINA v2.0: Mining interactome modules. Nucleic Acids Res. 2012;40(D1):862–5.

41. Barabási A-L, Gulbahce N, Loscalzo J. Network medicine: a network-based approach to human disease. Nat Rev Genet. 2011;12(1):56–68.

42. Nepusz T, Yu H, Paccanaro A. Detecting overlapping protein complexes in protein-protein interaction networks. Nature Methods. 2012. p. 471–2.

43. Reactome. Reactome - a curated knowledgebase of biological pathways. Genome Biology. 2008. p. R39.

44. Croft D, Mundo AF, Haw R, Milacic M, Weiser J, Wu G, et al. The Reactome pathway knowledgebase. Nucleic Acids Res. 2014;42(Database issue):D472–7.

45. Clemmons DR, Busby WH, Garmong A, Schultz DR, Howell DS, Altman RD, et al. Inhibition of Insulin-Like Growth Factor Binding Protein 5 Proteolysis in Articular Cartilage and Joint Fluid Results in Enhanced Concentrations of Insulin-Like Growth Factor 1 and Is Associated With Improved Osteoarthritis. Arthritis Rheum. 2002;46(3):694–703.

46. Busby WH, Yocum SA, Rowland M, Kellner D, Lazerwith S, Sverdrup F, et al. Complement 1s is the Serine Protease that Cleaves IGFBP-5 in Human Osteoarthritic Joint Fluid. Osteoarthr Cartil. 2009;17(4):547–55.

47. van Osch GJVM, van den Berg WB, Hunziker EB, Hauselmann HJ. Differential effects of IGF-1 and TGF beta-2 on the assembly of proteoglycans in pericellular and territorial matrix by cultured bovine articular chondrocytes. Osteoarthr Cartil. 1998;6:187–95.

48. Wang Q, Rozelle AL, Lepus CM, Scanzello CR, Song JJ, Larsen DM, et al. Identification of a central role for complement in osteoarthritis. Nat Med. 2012;17(12):1674–9.

49. Murphy G. Matrix metalloproteinases and their inhibitors. Acta Orthop Scand Suppl. 1995;266:55–60.

50. Wakitani S, Imoto K, Yamamoto T, Saito M, Murata N, Yoneda M. Human autologous culture expanded bone marrow mesenchymal cell transplantation for repair of cartilage defects in osteoarthritic knees. Osteoarthr Cartil. 2002;10(3):199–206.

51. Wakitant S, Okabe T, Horibe S, Mitsuoka T, Saito M, Koyama T, et al. Safety of autologous bone marrow-derived mesenchymal stem cell transplantation for cartilage repair in 41 patients with 45 joints followed for up to 11 years and 5 months. J Tissue Eng Regen Med. 2011;5(2):146–50.

52. Gruys E, Toussaint MJM, Niewold TA, Koopmans SJ. Acute phase reaction and acute phase proteins. J Zheihang Univ Sci B. 2005;6(11):1045–56.

53. Guney E, Menchem J, Vidal M, Barabasi A-L. Network-based in silico drug efficacy screening. Nat Commun. 2016;7:10331.

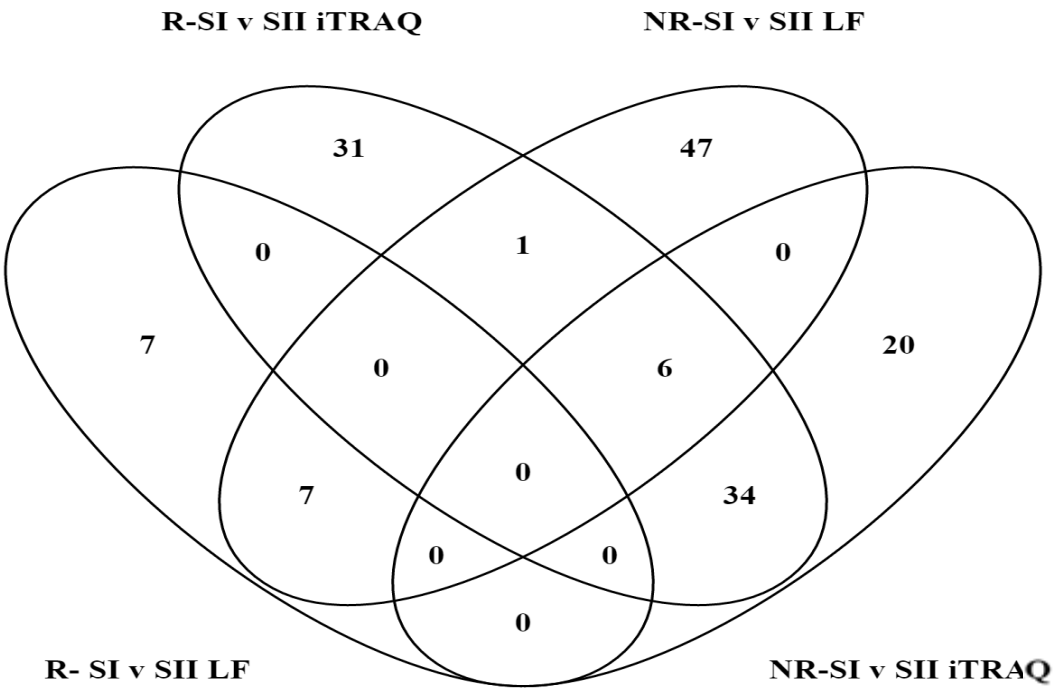


Figure 1: Venn-Diagrams representing the proteins identified using isobaric tags for relative and absolute quantitation (iTRAQ) proteomics and label-free quantitation (LF) proteomics which were differentially abundant (≥2.0 FC; p≤ 0.05) in the SF at Stage I (SI) compared to Stage II (SII) in responders (R) compared to non-responders (NR) to ACI.

742

743

744

745

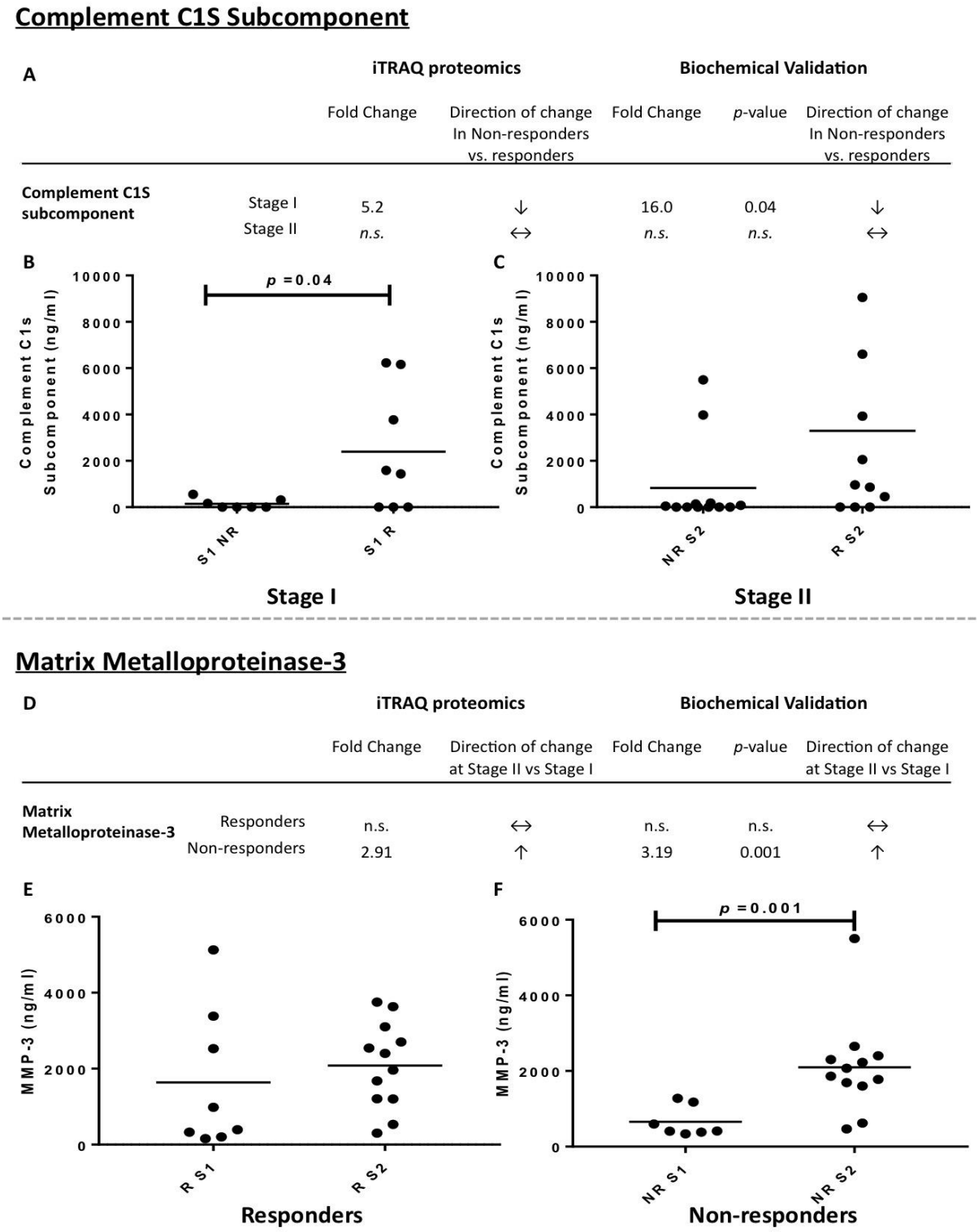


Figure 2: Biochemical validation of differentially abundant proteins identified using isobaric tagging for relative and absolute quantitation (iTRAQ) proteomics. (**A**) and (**D**) demonstrate the differential abundance of Complement C1S subcomponent and Matrix metalloproteinase-2, respectively as measured by iTRAQ mass-spectrometry and by biochemical ELISA. Quantitative ELISA confirmed that (**B**) Complement C1S subcomponent is significantly decreased in the synovial fluid (SF) of non-responders (NR) compared to responders (R) to Autologous Chondrocyte Implantation (ACI) prior to cartilage harvest (Stage I; S1; p =0.04; Student's t-test)(**C**) but was not significantly differentially abundant prior to chondrocyte implantation (Stage II; S2). (**E**) Matrix metalloproteinase-3 (MMP3) is not differentially abundant in response to cartilage harvest in ACI responders (**F**) but was biochemically confirmed to be differentially abundant in the SF of non-responders between Stages I and II of the ACI procedure (p =0.001; Student's t-test).

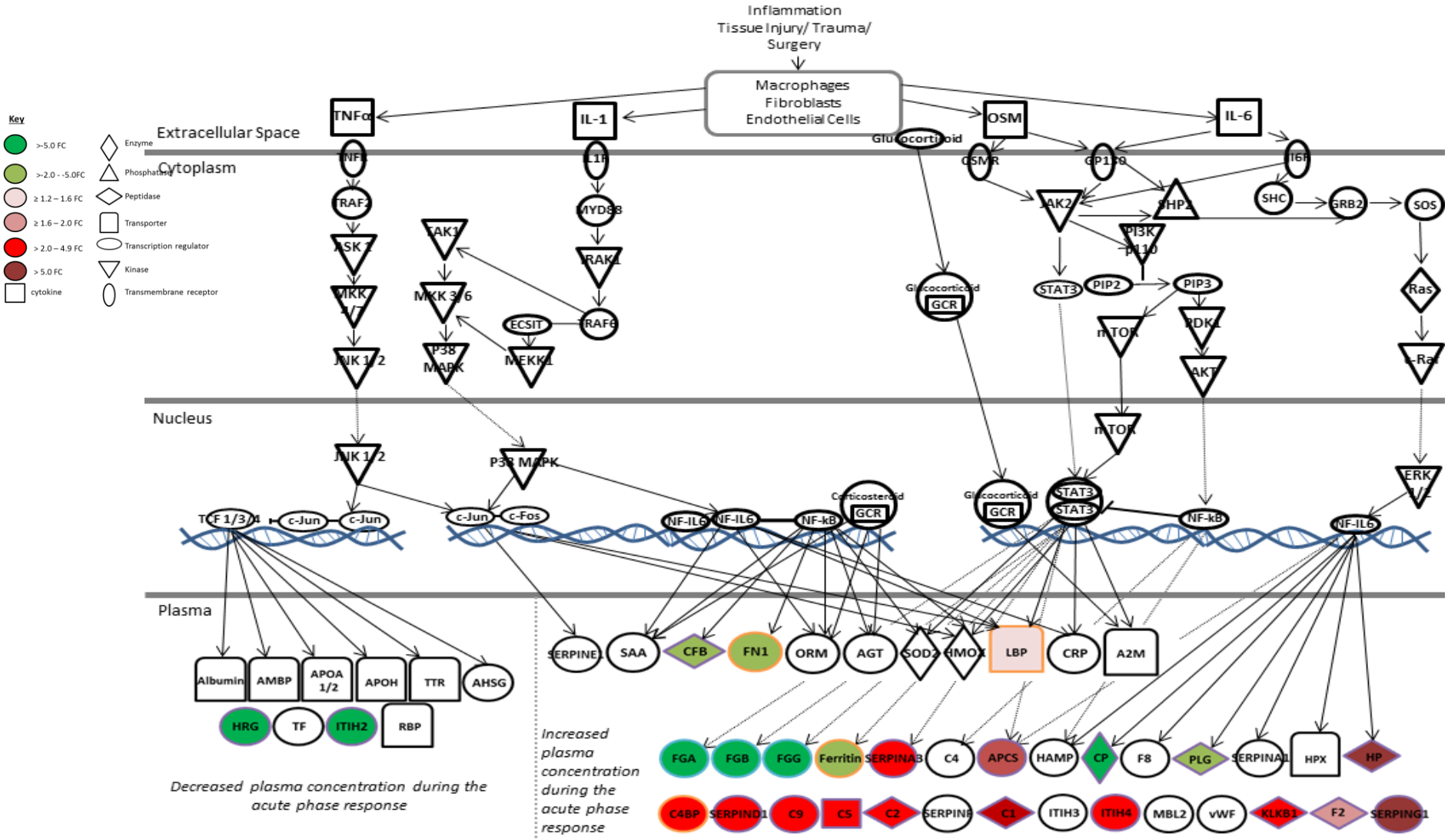
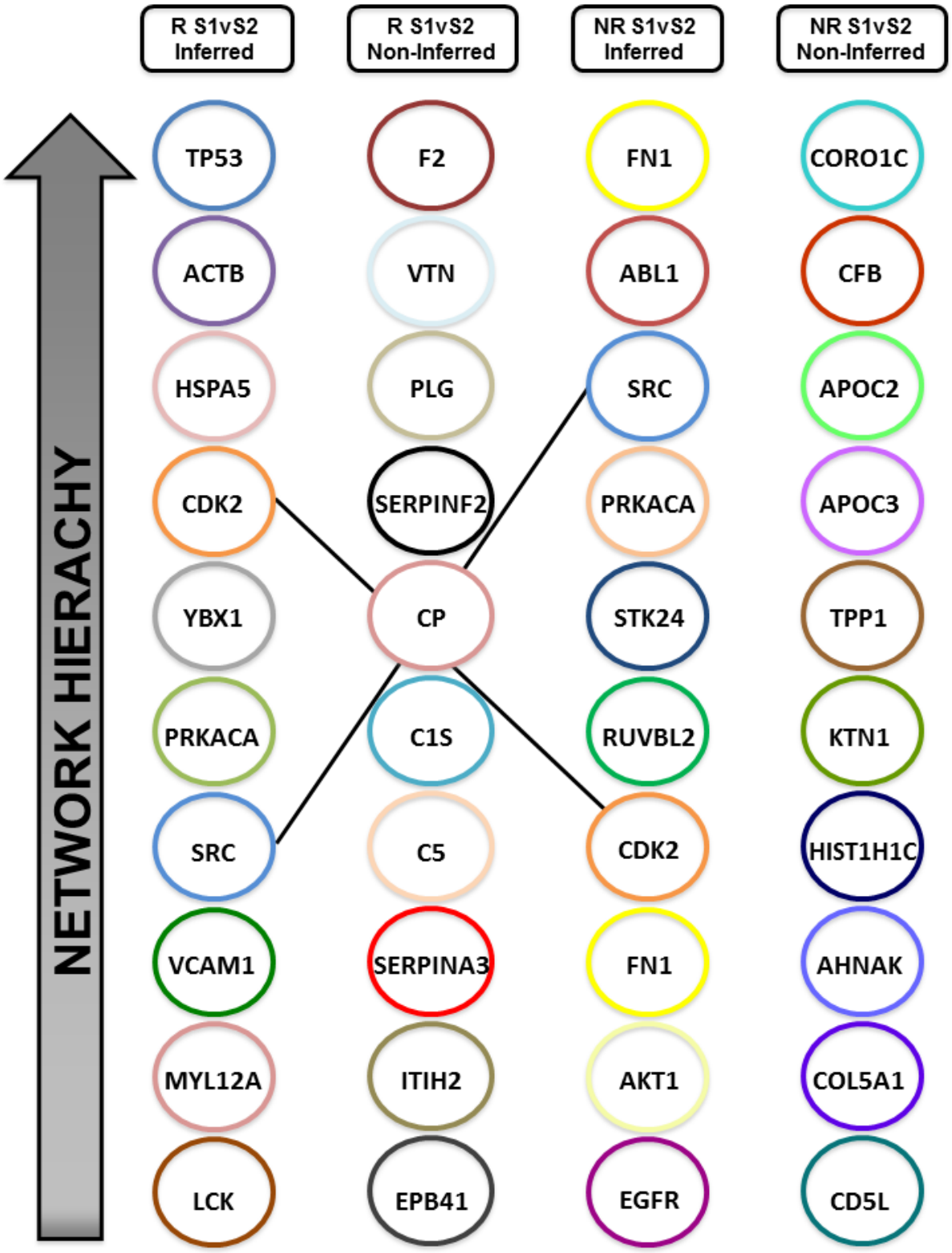


Figure 3: Proteins of Acute Phase Signalling at Stage II compared to Stage I in non-responders to Autologous Chondrocyte Implantation (ACI). Several synovial fluid proteins that are downstream of acute phase response signalling were differentially abundant between Stages I and II of ACI. Proteins edged in purple, orange and blue were identified using iTRAQ nLC-MS/MS, LF LC-MS/MS or by both techniques, respectively. (*Adapted from Ingenuity*).



772

773 **Figure 4: The ModuLand algorithm was applied to inferred and non-inferred interactome networks of**
774 **differentially abundant proteins (±1.2FC; p≤0.05) between Stages I and II of ACI in clinical responders and**
775 **non-responders.** Modules were identified from both non-inferred (protein changes identified from proteomic
776 analysis only) and inferred (identified protein changes and inferred proteins interactions) networks and are
777 ranked based on their hierarchical network connectivity.